Local Electric Fields Dictate Function: The Different Product Selectivity Observed for Fatty Acids Oxidation by Two Deceptively Very Similar P450-Peroxygenases OleT and BSβ

Shalini Yadav,[1] Sason Shaik*,[2] Shakir Ali Sidhdhiqui[1,] Surajit Kalita[1], and Kshatresh Dutta Dubey*[1,3]

1. Department of Chemistry, School of Natural Science, Shiv Nadar University Delhi-NCR, NH91 Tehsil Dadri, Greater Noida, Uttar Pradesh 201314, India

2. Institute of Chemistry, The Hebrew University, Edmond. J. Safra Campus, Givat Ram, Jerusalem 9190400, Israel. sason@yfaat.ch.huji.ac.il ORCID 0000-0001-7643-9421,

3. Center for Informatics, Department of Chemistry, School of Natural Science, Shiv Nadar University Delhi-NCR, NH91 Tehsil Dadri, Greater Noida, Uttar Pradesh 201314, India

   kshatresh.dubey@snu.edu.in [0000-0001-8865-7602]

Abstract:

Cytochrome P450 peroxygenases use hydrogen peroxide to hydroxylate long-chain fatty acids by bypassing the use of O_2 and a redox partner. Among the peroxygenases, *P450$_{OleT}$ uniquely performs decarboxylation of fatty acids and production of terminal olefins*. This root taken by P450$_{OleT}$ is intriguing, and its importance is augmented by the practical importance of olefin production. As such, this mechanistic choice merits elucidation. To address this puzzle we use hybrid QM/MM calculations and MD simulations for the OleT enzyme as well as for the structurally analogous enzyme, P450$_{BS\beta}$. The study of P450$_{OleT}$ reveals that the protonated His85 in the wild-type P450$_{OleT}$, *plays a crucial role in steering decarboxylation activity by stabilizing the corresponding hydroxoiron(IV) intermediate (Cpd II)*. In contrast, for P450$_{BS\beta}$ in which Q85 replaces H85, *the respective Cpd II species is unstable and it reacts readily with the substrate radical by rebound, producing hydroxylation products*. As we show, this single-site difference creates in P450$_{OleT}$ a local electric field (LEF), which is significantly higher than that in P450$_{BS\beta}$. *In turn, these LEF differences are responsible for the different stabilities of the respective Cpd II/radical intermediates, and hence for different functions of the two enzymes*. P450$_{BS\beta}$ uses the common rebound mechanism and leads to hydroxylation, whereas P450$_{OleT}$ proceeds via decarboxylation and generates terminal olefins. Olefin production projects the power of a single residue to alter the LEF and the enzyme's function!

1. Introduction

Cytochromes P450 are nature's efficient nanomachines that catalyze a plethora of reactions in all organisms.[1] These enzymes are phenomenally versatile; they are responsible for detoxifications of foreign compounds (so-called Xenobiotics), and among other processes can perform e.g. sulfoxidation, C-H/C=C activations, carbene- and nitrile-insertion, as well as C-C bond making and cleavage.[2-4] Such a synthetic versatility makes P450 a favorite model for bio-engineering and evolutionary studies.[5,6] The paradigmatic catalytic cycle of a P450 starts by the entry of a substrate, which triggers concomitant delivery of two electrons (from a reducing partner) and two protons, finally leading to the formation of the porphyrin-radical-cation oxoiron species, so-called Compound I (Cpd I) which in turn oxidizes the substrate.[7] This mechanistic paradigm has prevailed until the discovery of P450 peroxygenases which utilize hydrogen peroxide to shunt the catalytic cycle, and thereby bypasses the necessity of reducing partner by.[8]

The crystal structures of two P450 peroxygenases, $P450_{SP\alpha}$ and $P450_{BS\beta}$, reveal some interesting architectural features, related to substrate-binding.[9,10] In contrast to $P450_{BM3}$, which is a bacterial enzyme that catalyzes the hydroxylation of the tails of long-chain fatty acids, in the related peroxygenases $P450_{SP\alpha}$ and $P450_{BS\beta}$, the fatty acids enter the heme site via their carboxylate heads, and undergo hydroxylation, respectively in the α- or β- positions (to the carboxylate heads).

The unique structural aspect of peroxygenases is the interaction of the carboxylate group of the fatty acids with an Arginine residue close to the heme site. This feature bypasses the necessity of acid-alcohol pair for proton deliveries as in common P450s.[11] These special features of P450 peroxygenases indicate that this group of enzymes might have been bioengineered by Nature during primeval conditions which necessitate bypassing dioxygen requirement for the catalytic cycle.[12,13] Despite these differences, peroxygenases catalyze fatty acid oxidation, by use

of the same active oxidant, Cpd I, as in common P450s.[14,15] In the past, we have shown how the

H_2O_2 shunting generates Cpd I species in P450$_{SP\alpha}$ and P450$_{BS\beta}$,[16] which perform C-H

hydroxylations of fatty acids.[16]

So what is so special about peroxygenases? The discovery of P450$_{OleT}$,[17] has created a new

and challenging function. Thus, in P450$_{OleT}$ *Cpd I leads to preferential decarboxylation over*

hydroxylation, unlike the other peroxygenases. This puzzle is reinforced by the fact that P450$_{OleT}$,

P450$_{SP\alpha}$ and P450$_{BS\beta}$ exhibit *similar active site residues and substrate binding sites, and yet their*

function is different. In fact, as shown by Figure 1, P450$_{OleT}$ and P450$_{BS\beta}$ are extremely close with

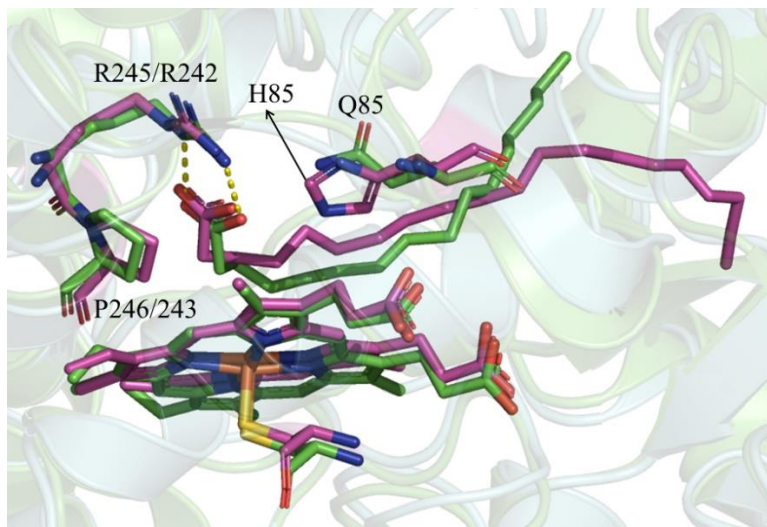an RMSD (root mean square deviations) of merely 0.99Å![8]

Figure 1. Superposition of P450$_{BS\beta}$ (green) and P450$_{OleT}$ (pink) crystal structures. The residue number are noted respectively in the order P450$_{OleT}$ / P450$_{BS\beta}$. Note that with the exception of H85(P450$_{OleT}$) and Q85(P450$_{BS\beta}$), all other residues occupy the same positions in the two enzymes.

The key difference, between these enzymes is a single residue adjacent to the heme; it is glutamine

(Q85) in P450$_{SP\alpha}$ and P450$_{BS\beta}$, whereas in P450$_{OleT}$ (cf. Fig 1) this residue is Histidine (H85).

Logically, this Histidine residue should be a plausible proton donor that enables the formation of

Cpd I from H_2O_2.[17b] This is further supported by the fact that a Q85H (Q is Glutamine (Gln)) mutation in P450$_{BSβ}$ increases the decarboxylation activity over hydroxylation in the so-mutated P450$_{BSβ}$.[18] Suggestions that the substrate identity is somehow involved in changing the mechanism were ruled out by Reetz et al.[6d] As such, the primary mechanistic puzzle which is targeted here, is: *how does the presence or absence of H85 diversify the decarboxylation and hydroxylation activities in the WT P450$_{BSβ}$ and WT P450$_{OleT}$ P450 enzymes*? These issues are addressed herein by means of a combination of MD simulations and QM/MM studies of the P450$_{OleT}$ and P450$_{BSβ}$ enzymes. As shall be seen, the study uncovers the nimble ability of evolution to bioengineer reactions using a mutation of a single protein residue, which affects its local electric field (LEF)[19] and hence its function.

Scheme 1 shows the plausible mechanisms of decarboxylation vs hydroxylation, which have been considered in peroxygenases. As seen, the hydroxylation in P450$_{BSβ}$ performs the standard hydrogen atom transfer (HAT) by Cpd I which is converted to Cpd II, followed by a rebound step to yield a hydroxylated product. However, three alternative pathways were proposed for the decarboxylation mechanism in P450$_{OleT}$ (Scheme 1).[8] Pathway 1 involves a single electron transfer from the substrate to Cpd I followed by HAT to the so formed porphyrin-oxoiron(IV), Cpd II. However, this pathway was ruled out by the absence of OleT accumulation for reactions with a protonated long-chain fatty acid, and at the same time, the reaction of OleT exhibited a significant substrate deuterium kinetic isotope effect (KIE).[12] Pathway 2, in Scheme 1, was ruled out too due to spectroscopic observation of a metabolite with a Soret band maximum at 426 nm and hyperporphyrin (i.e. a species having an additional band between 300 nm and 426 nm) character, corresponding closely to Cpd II.[14,20]
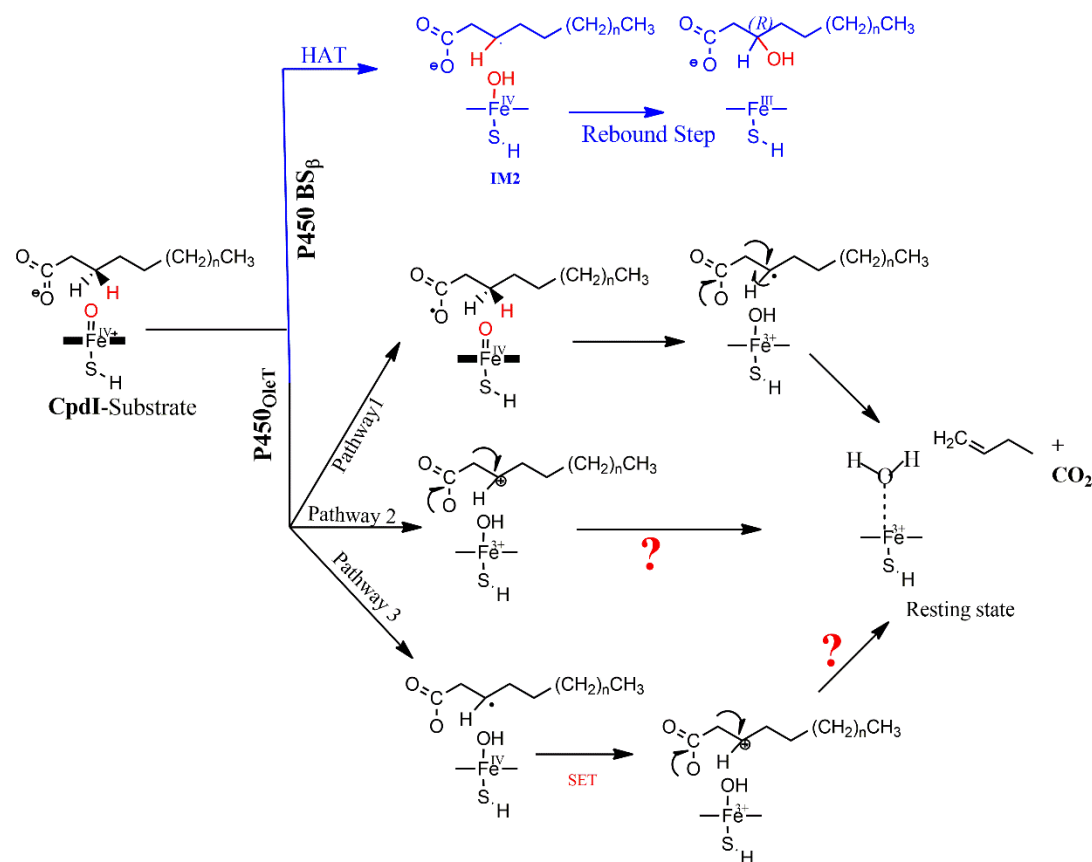
Scheme 1. Plausible mechanisms of hydroxylation by $P450_{BS\beta}$ and decarboxylation by $P450_{OleT}$. The question marks in pathways 2 and 3 refer to the mechanistic puzzle of the nature of the proton sources in these mechanistic pathways. This puzzle is targeted in this study.

Pathway 3 is based on the observation that the protonated Cpd II (the hydroxoiron(IV) intermediate), which was observed in the decay kinetics, has a very slow conversion to the resting state, thus implying a relatively stable Cpd II which accepts a single electron transfer (SET) from the substrate, which in turn, loses CO_2 and form the terminal olefin.[21] However, this pathway does not explain why this mechanism should be utilized only by $P450_{OleT}$ but not by the very similar $P450_{BS\beta}$ enzyme. Thus, mechanistic studies did not resolve the origins of *decarboxylation in P450_{OleT}, and the identity of the factors which might drive P450_{OleT} to adopt a different mechanistic pathway vis-à-vis P450_{SP\alpha} and P450_{BS\beta}.* The industrial uses of olefins lend importance to this specific propensity of OleT to produce olefins.

To resolve the issue, we applied herein MD simulations followed by QM/MM calculations, for WT P450$_{OleT}$ and P450$_{BS\beta}$. These calculations lead to a *revised mechanism through which OleT performs decarboxylation of fatty acids, due to the assistance of a protonated His85*. This residue creates a local electric field (LEF), which is larger than the one created by Q85 and ultimately dictates the product selectivities of P450$_{OleT}$ vs. P450$_{BS\beta}$.

2. Computational Methods and Details

We used Molecular Dynamics (MD) simulations for the conformational analysis and hybrid QM/MM calculations for the reaction mechanism; the technical details are discussed below:

2.1 System Setup:

The study focuses on the following enzymes; P450$_{OleT}$ and P450$_{BS\beta}$. The initial structures for these systems were taken from the crystal structure with PDB accession codes: 4L40,[17] and 1IZO[21] respectively. Missing hydrogens and other heavy atoms were added by Leap module of Amber 20. The parameters for the substrate Palmitoloeic and Icosanoic acids were prepared using antechamber module of Amber 20 for GAFF2 parameters. The partial atomic charges and missing parameters for both substrates were obtained from the RESP[22,23] charge fitting method for the QM optimized geometry at the HF/6-31G* level of theory. The force field for the heme moiety was taken from already published parameters for the Cpd I state.[24] Na+ ions were added into the protein surface to neutralize the total charge of the system. Finally, the resulting system was solvated in a rectangular box of TIP3P[25] waters extending up to a minimum cutoff of 10 Å from the protein boundary.

2.2 MD Simulations:

After proper parameterizations and setup, the resulting system's geometries were minimized (5000 steps for steepest descent and 10,000 steps for conjugate gradient) to remove the poor contacts and relax the system. The systems were then gently annealed from 10 to 300 K under canonical ensemble for 50 ps with a weak restraint of 5 kcal/mol/Å^2. Subsequently, the systems were maintained for 1 ns of density equilibration under isothermal–isobaric ensemble at the target temperature of 300K and the target pressure of 1.0 atm using a Langevin-thermostat[26] and a Berendsen barostat[27] with a collision frequency of 2 ps and a pressure relaxation time of 1 ps, and with a weak restraint of 1 kcal/mol/Å^2. Thereafter, we removed all restraints applied during heating and density dynamics and further equilibrated the systems for ~3 ns to obtain well-settled pressure and temperature for conformational and thermodynamical analyses. This was followed by a productive MD run of 100 ns for each system. We used three different replicas starting from different initial velocities each for 100 ns. Therefore, we performed a total of 300 ns simulations including all three replicas for each system. During all MD simulations, the covalent bonds containing hydrogen were constrained using SHAKE[28], and a particle mesh Ewald (PME)[29] was used to treat long-range electrostatic interactions. We used an integration step of 2 fs during the entire simulation. All MD simulations were performed with the GPU version of the Amber 20 package.[30]

The analyses of the trajectories were based on the most populated trajectories (100ns) from the MD simulations of all three replicas ($100ns \times 3$), which provides an accurate means to represent the statistically most significant structures during MD sampling. For doing so, firstly, we seamed all replicas together using the Cpptraj module of Amber, and thereafter, we performed clustering of MD trajectories using the hieragglo algorithm implemented in Cpptraj.

2.3. QM/MM Calculations:

The reaction mechanism was investigated by use of the QM/MM calculations for the representative snapshots of the most populated trajectories obtained by clustering the MD trajectories for each system. This choice is statistically more rigorous than the usage of stochastically chosen snapshots for QM/MM calculations.

QM regions included truncated heme-porphyrin, truncated substrates, and R242/245, H85/Q85 and two water molecules, the coordinates for the QM region can be found in SI. All protein residues and water molecules within 8Å of the Heme were included in the 'active region' of the QM/MM calculations. The atoms in the 'active region' interact with the QM atoms through electrostatic and van der Waals interactions and the corresponding polarization effects were considered in the subsequent QM/MM calculations. All QM/MM calculations were performed with ChemShell,[31,32] by combining Turbomole[33] for the QM part, and DL_POLY[34] for the MM part. The MM region was described using the Amber ff14SB force field,[35] and the electronic embedding scheme was used to account for the polarizing effect of the enzyme environment on the QM region. The QM/MM boundary was treated using hydrogen link atoms with the charge-shift model.[32]

During QM/MM geometry optimizations, the QM region was treated using the hybrid UB3LYP[36] functional with two basis sets, B1 and B2, where B1 stands def2-SVP* and B2 is for def2-TZVP. For geometry optimization and frequency calculations, we used B1 basis set. The energy was further corrected using B3LYP-D3[37] functional. All of the QM/MM transition states (TSs) were located by relaxed potential energy surface (PES) scans followed by full TS optimizations using the P-RFO optimizer[38] implemented in the HDLC code. The energies were further corrected with the large all-electron def2-TZVP basis set. The zero-point energies (ZPEs) were calculated for all species, and the respective final energies are reported as UB3LYP/B2+ZPE

data. For Cpd I-mediated reactions, the S = 3/2 and S = 1/2 states generally exhibit similar reactivities,[4b] so no added information (other the well know two-state reactivity) is provided by studying both states. Hence, all of the reactions of Cpd I were studied in the S = 1/2 state.

2.4. Determination of the LEFs of the Enzymatic Environments

The local electric field (LEF) of both enzymes, P450$_{OleT}$ and P450$_{BS\beta}$ were calculated by means of the in-house program TITAN.[39] To qualify the effect of the active site residues only, we stripped all protein and water molecules beyond 3Å of the Heme and substrate, such that both enzymes (the and Heme-Cys/substrate systems of P450$_{OleT}$ / P450$_{BS\beta}$) were subjected to the LEFs exerted respectively by H85/Q85, R245/R242, P246/P243, and two water molecules. Therefore, it is clear that the change in the LEF will be mainly due to the difference by H85 vs. Q85. The LEF was calculated along Fe-O axis, projecting on the O atoms of both enzymes. Similarly, the dipole moments (μ) for the QM/MM optimized bare heme-Cysteine/substrate systems were determined with Turbomole 7.31.[40]

The stabilization energies were calculated by use of equation 1:[41]

$$\Delta E_{LEF} = 4.8(LEF)(\mu)cos\theta \qquad (1)$$

ΔE is the interaction energy (kcal/mol), between the LEF (in V/Å) of the protein environment and the dipole moment μ (in Debye) of the heme species, while θ is the angle between the dipole moment and LEF vectors. These calculations were done for the RC,TS (Cpd I/substrate) and the Cpd II/substrate-radical species for each reaction. The GAUSSIAN convention was used for the positive directions of the two dipoles.[41]

3. Results

3.1 MD Simulations of P450$_{BS\beta}$ and P450$_{OleT}$: The MD simulations for each enzyme/substrate set included 100 ns with three replicas, as well as usage of random velocities. The calculations were performed for P450$_{OleT}$ and P450$_{BS\beta}$ reacting with their corresponding substrates; icosanoic acid [$CH_3(CH_2)_{18}COOH$] and palmitic acid [$CH_3(CH_2)_{14}COOH$], respectively.

The MD simulations in both enzymes revealed very stable substrate binding, via strong salt-bridge interactions of the guanidine moiety of Arginine (R, cf. Figure. 1) with the carboxylate ends of either substrate. In addition, as shown in Figure 2 both simulations reveal the presence of a consistent water chain mediated by a heme-propionate (also see supplementary video VS1 and VS2 for water channel). The presence of a water molecule is considered to have an important function in peroxygenases.[16] Therefore, our simulations for two important members of peroxygenase family show that peroxygenases may have evolved, during the primitive anaerobic conditions, in such a way that they preserved the proton transfer route (via heme propionate) while bypassing the other route (acid-alcohol pair).

Though the two-water channels are rather similar in both peroxygenases, *we found a significant difference in the behaviors of the substrates.* In P450$_{OleT}$ (cf. Fig. 2a), the substrate (abbreviated as DCR) interacts with His85 via a bridging water W1, which causes the substrate to recline towards the His85. The first oxygen O1 of the carboxylate group in P450$_{OleT}$ forms a bidentate interaction with the guanidine groups of the Arginine residue (R245), while the second oxygen, O2, interacts with the two water molecules W1 and W2 (cf. Fig. 2a). Interestingly, an equivalent interaction between the substrate (PAM) and Gln 85 in P450$_{BS\beta}$ is completely absent. In fact, in P450$_{BS\beta}$ the substrate stays far away (> 5Å) from the Gln 85 (see Figures 2b and 2c). Thus, in P450$_{BS\beta}$ (cf. Fig. 2b), the carboxylate oxygen, O1, interacts with the guanidine nitrogen NH1 (of R242), while O2 forms a bidentate interaction with W2 and with the guanidine nitrogen

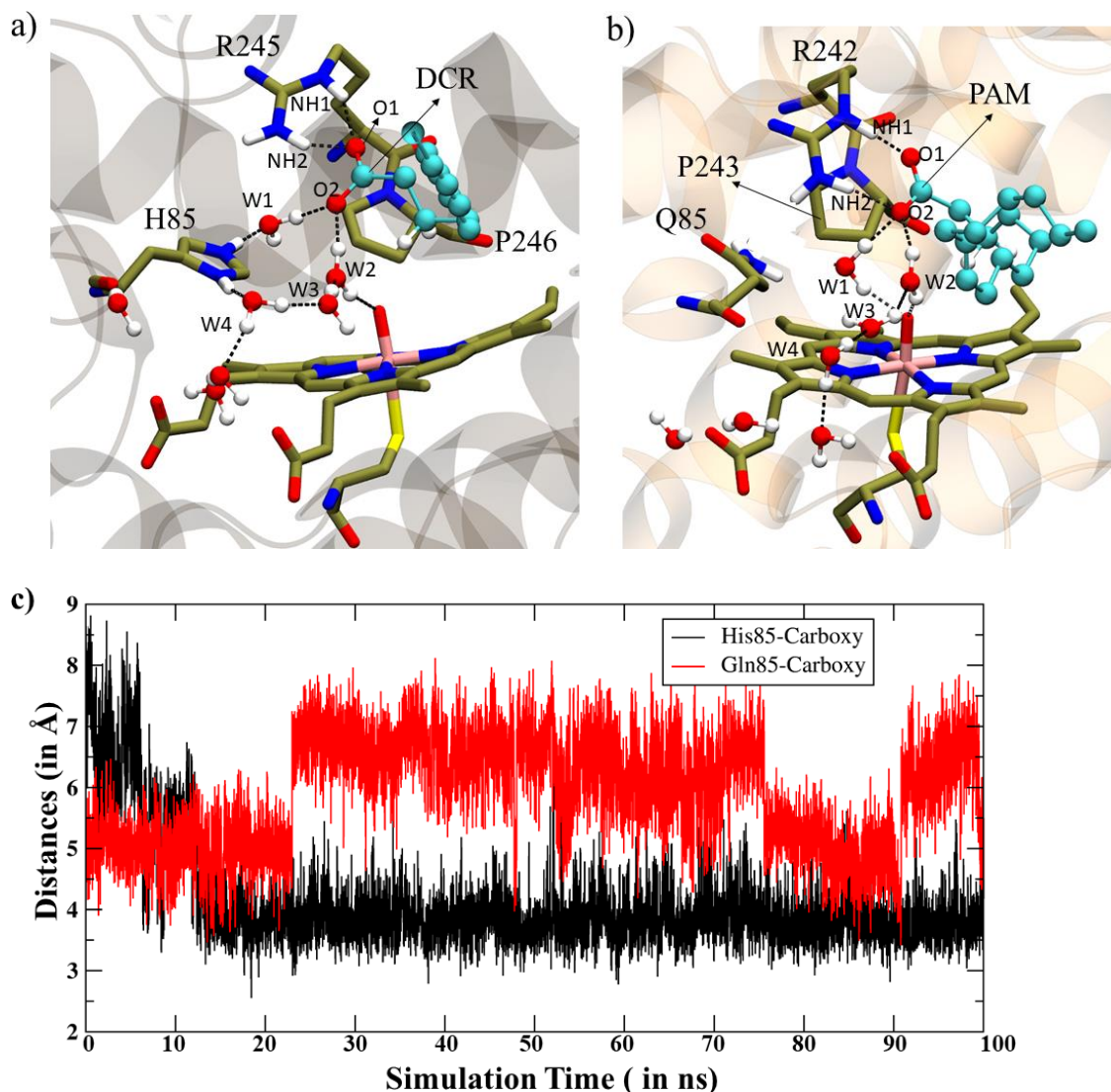(NH$_2$) of R242. Figure 2c compares the time evolution of these key interactions in P450$_{OleT}$ (black) and P450$_{BSβ}$ (red).

Figure 2. A comparison of substrate orientations in two P450 peroxygenases a) P450$_{OleT}$; here DCR represents the Icosanoic acid [CH$_3$(CH$_2$)$_{18}$COOH], b) P450$_{BSβ}$; here PAM represents palmitoleic acid, CH$_3$(CH$_2$)$_{14}$COOH. c) The distances during MD simulations between the carboxyl groups and His85 in P450$_{OleT}$ (black), and with Gln85 in P450$_{BSβ}$ (red). Note that the interaction of the carboxylate group of DCR with His85 (in a) is mediated by W1 which is lacking in P450$_{BSβ}$ (in b).

In summary, the MD simulations of P450$_{OleT}$ vs. P450$_{BSβ}$ show that a single site-difference in residues 85, of H vs. Q, changes the active site architecture in the two peroxygenases. The

manner by which this difference is translated to different functions of the peroxygenases is detailed in the next section with the help of hybrid QM/MM calculations.

3.2. Mechanism of decarboxylation vs hydroxylation by QM/MM calculations

Decarboxylation in P450$_{OleT}$: To study the mechanism of decarboxylation reaction in the WT P450$_{OleT}$ enzyme, we performed QMMM calculations for H atom transfer (HAT) from the prochiral CH$_2$ group in the DCR substrate, employing a representative snapshot from the most populated structure of the MD simulations. The optimized structure of the reactants' complex (RC) is shown in Figure 3. The Pro-*R* hydrogen of the β position is positioned at 2.6 Å and is closer to the oxo moiety of Cpd I than the Pro-*S* hydrogen at 3.3Å. Moreover, in the optimized geometry of the reactant cluster (RC) the O---H bond length of water W1 is 1.08Å, namely, longer than the pristine O-H bond length (0.98 Å) in the water molecules (cf. Figure 3). The potential energy profile for the doublet-spin state for Pro-*R* H abstraction (cf. Figure 3) exhibits an energy barrier of 9.6 kcal/mol, thus revealing a very facile H-abstraction from the β position. To assure the correct energy profile all the way to Cpd II, we also performed the calculations, using the ChemShell routine Nudged Elastic Band, for the H-abstraction process and the energy profile was practically identical (9.9 kcal/mol).

The potential energy profile in Figure 3 shows also the follow-up steps all the way to CO$_2$ expulsion and olefin formation. The changes in orbital occupation during all the steps of Figure 3 are shown in Figure 4. Thus, the first H-abstraction step (RC→IM1) transforms Cpd I to Cpd II (Fe(IV)OH), leaving a •C$_\beta$H radical carrying a CO$_2^-$ group in an α-position. At the same time, W1 stabilizes IM1 and appears to blocks the rebound step (cf. Figure 3). The second step (IM1→IM2), is a single electron transfer (SET) process wherein an electron from the carboxylate group of the substrate shifts to Fe (IV)OH and forms Fe (III)OH$^-$. Concomitantly, CO$_2$ splits and generates a

13

terminal double bond across C_α—C_β of the substrate (cf. Figs 3 and 4). The energy barrier for this step is 12.4 kcal/mol relative to IM1 which is lower than the calculated rebound barrier at $\bullet C_\beta H$.[42] In the final step, the protonated His85 relays a proton to the negatively charged IM2 via W1, thereby forming the aqua ferric low spin resting state and resets the catalytic cycle for the next turnover.
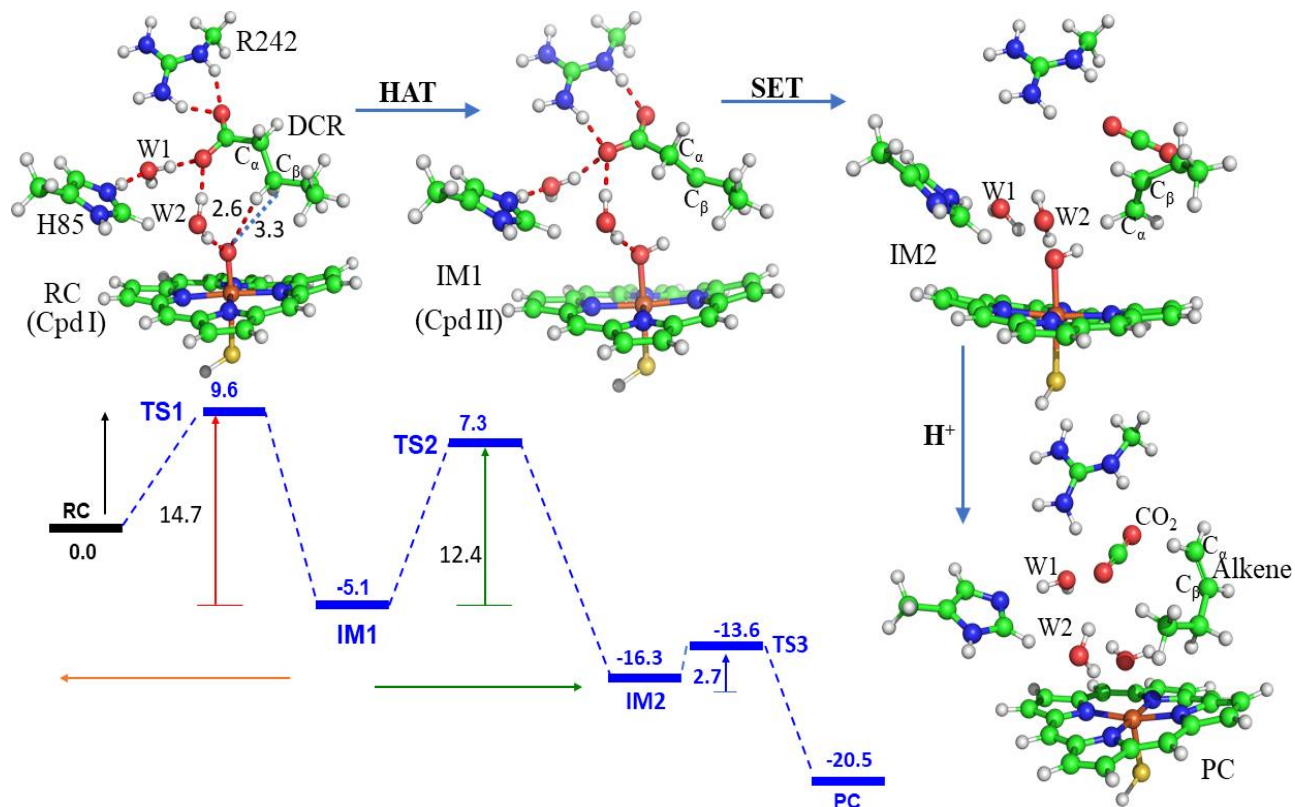

Figure 3. QM/MM/B3LYP-D3/def2-TZVP calculated mechanism for OleT oxidation of the substrate decadonoic acid. Key intermediates during the decarboxylation mechanism in OleT are shown by the ball and stick model. A corresponding energy profile and barriers for the process are shown in the schematic diagram at bottom of the figure. The reported energy values are in kcal/mol.
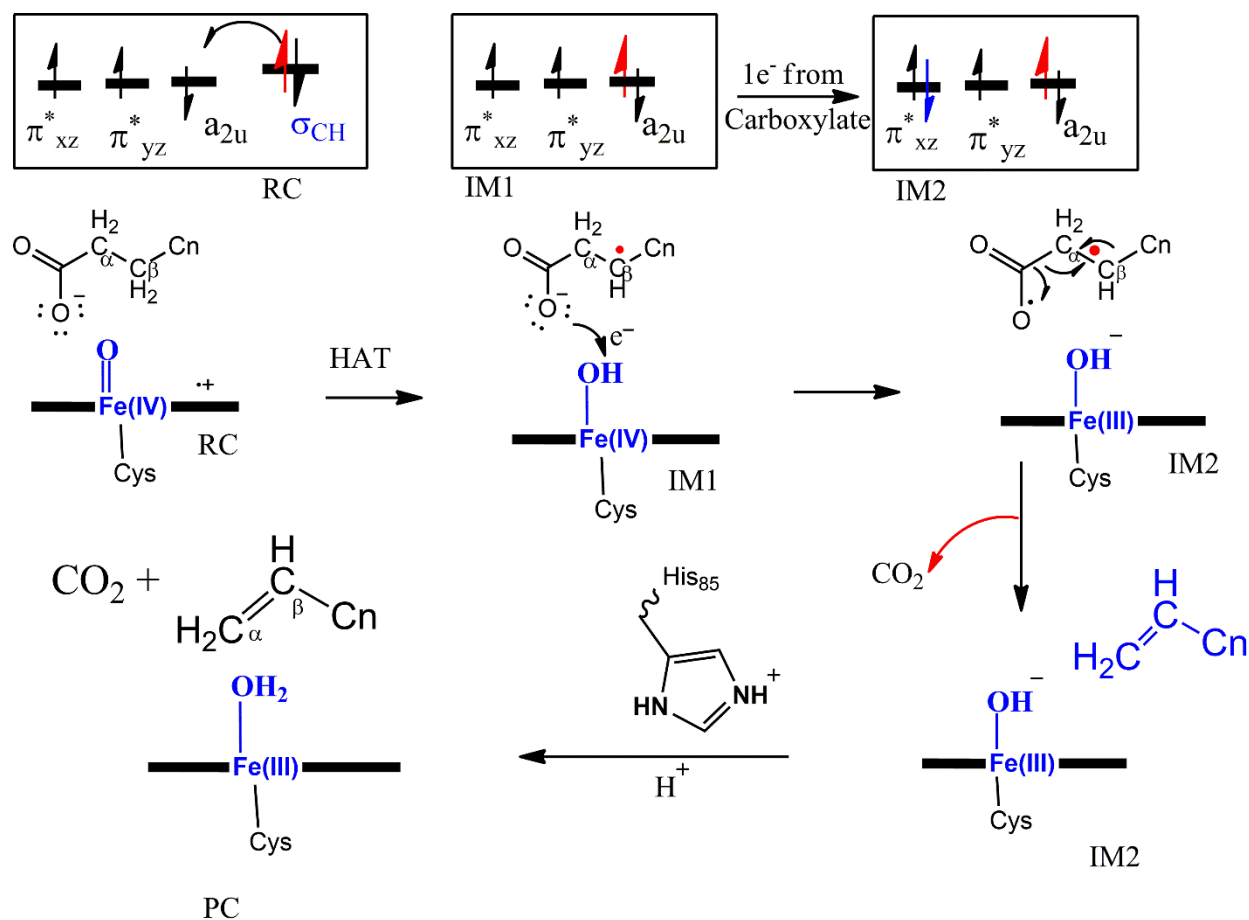
Figure 4. The orbital occupancy diagrams and the sequential steps during the mechanism of OleT decarboxylations in Figure 3. The first line of orbital diagrams describes the main changes of orbital occupancies, while below this line the mechanism is described sequentially.

The electronic structures of the intermediates in Figure 4 were confirmed by the spin densities and Mulliken charge analysis of intermediates (See SI). Interestingly, the intermediate IM1 i.e., Fe (IV)-OH/substrate-radical is exothermic by ~5 kcal/mol relative to Cpd I/substrate (cf. Figure 3), and the energy barrier for the back reaction from IM1→Cpd I is higher than the one for the forward reaction (IM1→IM2). *This observation qualitatively agrees with the experimental study by Grant et a[14] where they did not find any trace of Cpd I once Cpd I decays to IM1 (k_{-2} ~2 s^{-1}) and accumulation of intermediate IM1*. Another study by the same group shows that Cpd I,

decays very quickly (in 1 s), to a species that is indistinguishable from the ferric-low spin substrate-free form of the enzyme, and the time course for this decay *required two summed exponentials for adequate fitting, suggesting a complex decomposition process that may involve multiple steps.* Looking at the energy profile (c.f. Fig 3), we can see that the decay of Cpd I is an efficient process with a barrier of ~9.6 kcal/mol followed by a multistep process that involves Cpd I→IM1→IM2→product. *As such, our QM/MM calculations qualitatively agree with the conclusions of Grant et al.[14]*

It is apparent that our calculations provide compelling evidence that the peculiar mechanism of substrate decarboxylation in OleT is virtually a concerted process (barrier ~2 kcal/mol; cf. Figure 3) and the active site topology with its protonated His85 residue, is the root cause for the decarboxylation, whereas the common rebound process is blocked by an organized His85-water dyad (see W1,2 and H85 in IM1 cf. Figure 3).

Substrate Hydroxylation in P450$_{BS\beta}$: As we have seen above, *P450$_{OleT}$* does not proceed via a rebound mechanism, but prefers to accept one proton from His85, which occurs instantly and resets the cycle to the resting state. However, to substantiate the role of His85 that actually prevents the rebound step by associating a water molecule, we performed QMMM calculations for P450$_{BS\beta}$ which performs hydroxylation in the WT, where the His85 residue is naturally replaced by Glutamine (Q). The optimized reactant-complex geometry (RC) shows an organized water channel, however, unlike in P450$_{OleT}$, in P450$_{BS\beta}$ *the substrate stays far from the Q85 residue*. In fact, the carboxylate group of the substrate does not interact with Q85 by any means. The pro-*R* hydrogen at the C$_\beta$ terminal is the closest hydrogen to the FeO moiety.

The energy profile during the scanning is shown in Figure 5, which reveals firstly a H-abstraction step from the Pro-*R* beta hydrogen. The process takes place with a barrier of 21.2

kcal/mol leading to IM1 in an endothermic process. The spin densities of the iron-porphyrin (see SI) confirm IM1 as the Fe(IV)OH Cpd II intermediate. It is obvious that the H abstraction by Cpd I is endothermic, and as such, the unstable IM intermediate quickly follows the rebound mechanism to form the resting state. The QM/MM calculation for the P450$_{BS\beta}$ shows a similar energy profile to the one obtained in P450$_{SP\alpha}$ in a previous study.[16]
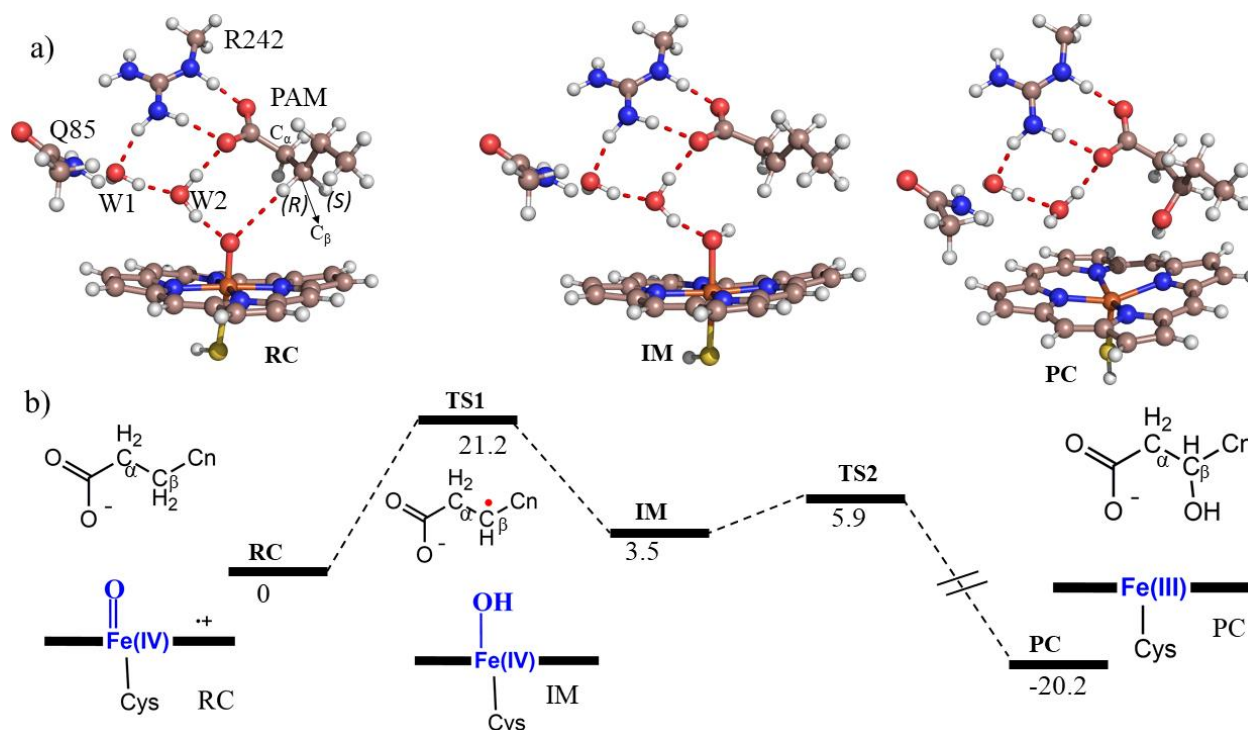

Figure 5. a) Different intermediates during the hydroxylation of Palmitic acids by P450$_{BS\beta}$, b) Energy profile during the PES scanning by QM/MM calculations at B3LYP-D3/def2-TZVP level of theory. In the absence of a proton donating group like His85, Cpd II follows the usual rebound step.

3.3. The Roles of Local Electric Fields (LEFs) in P450$_{OleT}$ and P450$_{BS\beta}$

As we have just seen, the stability of Cpd II/substrate-radical appears to control the decarboxylations vs hydroxylations reactions. As such, we proceeded to calculate the local electric fields (LEFs) and the dipole moments (μ), which are required in order to assess the stabilization

(cf. Eq. 1) of the RC, TS and IM1 species belonging to the respective reaction mechanisms of OleT

vs. BSβ (cf. Figs 3 and 5). To do so, we used the bare systems without residues, and applied the

LEFs due to the residues within a radius of 3Å around the species, using the in-house program

TITAN[39] (see method section for details). Figure 6a shows the conventions for positive vectors

along the Z-axis, while Figures 6b and 6c show the LEF and μ vectors, for the two enzymes, as

well as the angle between these vectors. It is apparent that for all the species, OleT has more

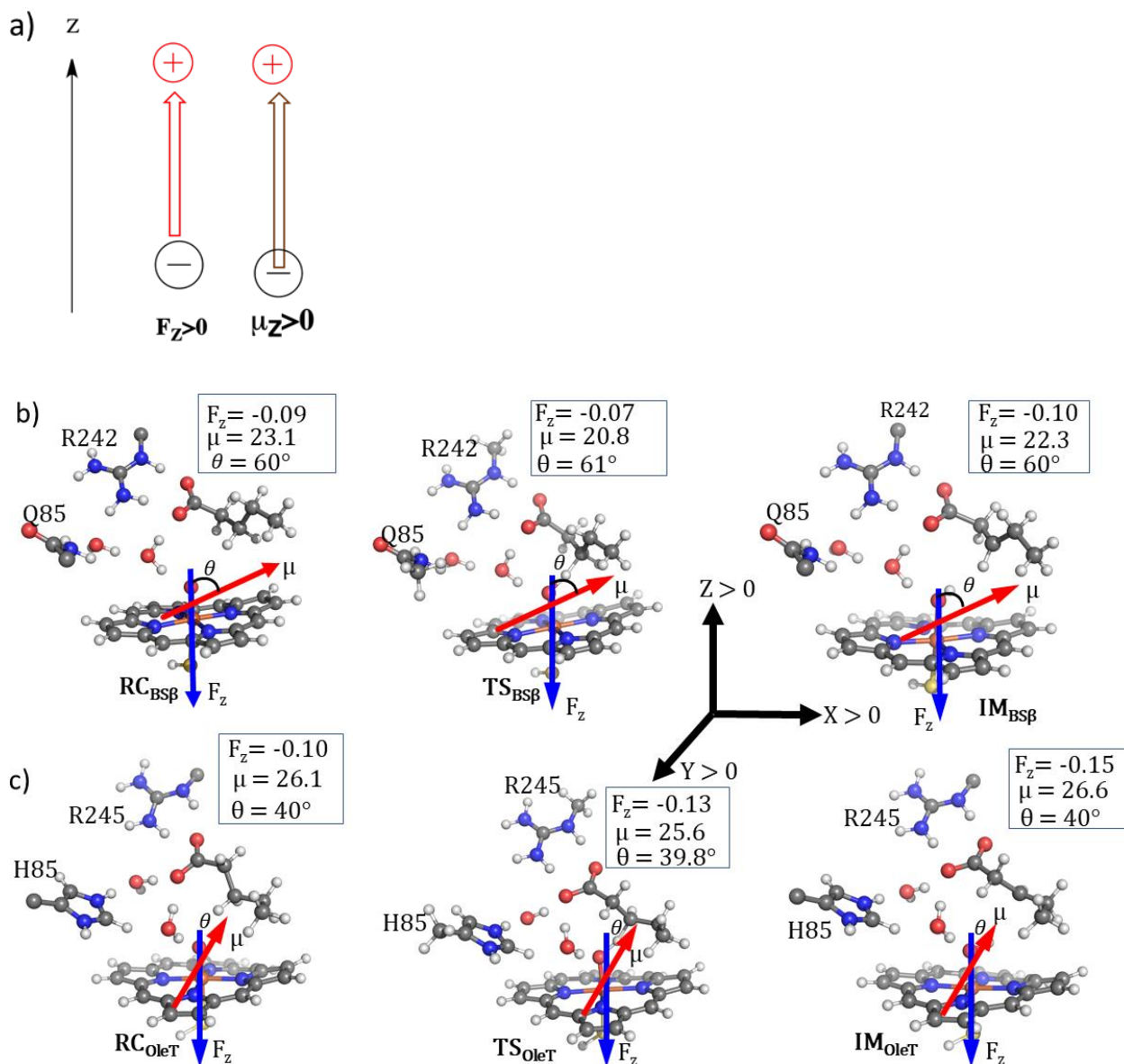favorable LEFs and dipole moments for each of the species.

Figure 6. (a) The GAUSSIAN convention for positive field and dipole moment vectors (F_Z and μ_Z) along the Z-axis. (b and c) The dipole moments μ (in Debye) and LEFs (F_Z in V/Å) along the Z-axis, for the RC, TS, and IM species of P450$_{BS\beta}$ and P450$_{OleT}$, respectively. LEFs were calculated in the native environment of both enzymes on the O atom of Fe-O.

Table 1 lists the corresponding stabilization energies of the species for the two enzymes by the corresponding LEFs. It is apparent that the LEF stabilizes the OleT species more effectively than it does for those of BSβ. The major LEF effect appears on Cpd II of the two enzymes. Comparison of the stabilization of the two Cpd II intermediates (IM) reveals that the LEF contributes 16.7 kcal/mol to the enhanced stability of Cpd II (OleT) compare with only 5.4 kcal/mol to Cpd II (BSβ). This difference (11.3 kcal/mol) is rather close to the same difference in the total QM/MM energy difference (8.4 kcal/mol) between the two species. (cf. Figs. 3 and 5).

Table 1. Interaction energies (in kcal/mol) due to the interaction of local electric field (F_Z) for reactant (RC), transition state (TS) and compound II intermediate (IM) for both enzymes. Note that a negative value indicates a relative stabilization during the reaction.

Enzyme	RC	TS	IM
P450$_{OleT}$	-10.61	-13.89	-16.66
P450$_{BS\beta}$	-4.83	-3.49	-5.35

It is apparent that the LEF contributes also to a lower barrier for H-abstraction in OleT vs. BSβ (~4.6 kcal/mol). However, the QM/MM-calculated energy-barrier difference is 10.6 kcal/mol. This suggests that effects other than the LEF contribute to the energy barrier differences between

the two enzymes. We propose that the more pronounced energy-barrier lowering in OleT vs. BSβ, is contributed by the relative thermodynamic driving forces of the corresponding H-abstraction steps. This step is endothermic in BSβ (cf. Fig. 5), due to the instability of the corresponding Cpd II species, whereas in OleT it is exothermic (cf. Fig. 3) due to the stability of Cpd II. Since the relative stabilities of the Cpd II species in the two enzymes are determined by the respective LEFs (cf. Figs 6b vs. 6c), the ensuing thermodynamic control of the H-abstraction barrier[43] is an indirect effect of the LEFs in the two enzymes.

Considering residues that are contained in the radius of 3Å of the bare QM species, it is apparent that the LEFs exerted by the two proteins on Cpd II differ only by protonated His 85 in OleT vs. Gln 85 in BSβ. As such, the LEF determines the enhanced stability of Cpd II (OleT) compared to Cpd II (BSβ), *and creates thereby function in these two enzymes*. As such, *Nature has performed ultimate engineering in P450$_{OleT}$ by substituting His85 to produce an electrostatic environment that can stabilize the Cpd II and thereby catalyze the decarboxylation reaction*. The absence of H85 in P450$_{BSβ}$ creates a less polar environment, and the enzyme follows the paradigmatic rebound mechanism with an unstable Cpd II/substrate-radical. This diversification is summarized in Scheme 2.
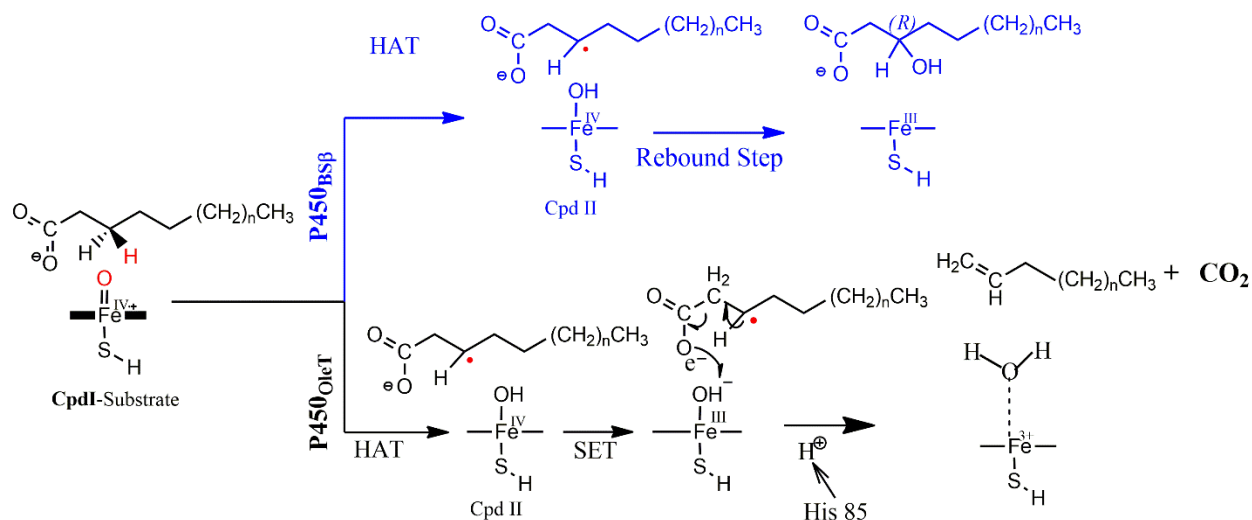
Scheme 2: The updated mechanism of hydroxylation (blue) vs decarboxylation (black). Note the source of a proton is His85 in WT P450$_{OleT}$.

4. Conclusions:

P450$_{OleT}$ and P450$_{BSβ}$ have virtually the same active sites save a single residue difference, H85 in P450$_{OleT}$ is Q85 in P450$_{BSβ}$. Nevertheless, the enzymes exhibit different functions; P450$_{OleT}$ converts fatty acids to terminal olefins, while P450$_{BSβ}$ hydroxylates the β C-H bond of the fatty acid. Using MD simulations and hybrid QM/MM calculations we demonstrated that *it is this single site difference that creates the functional difference by diversifying the corresponding LEFs of the active sites*. Thus, in P450$_{OleT}$, the strong LEF due to the presence of strategic protonated residue His85 stabilizes the respective Cpd II/substrate-radical species and diverts it to perform decarboxylation and produce terminal olefins. On the other hand, the presence of Q85 in P450$_{BSβ}$ creates a smaller LEF, which gives rise to an unstable Cpd II/substrate-radical species, which in turn proceeds by hydroxylation via the usual rebound mechanism.

Our finding that a polar environment in peroxygenase active site triggers decarboxylation reaction provides a lucid understanding of the function of P450$_{OleT}$ and may constitute a helpful lead for bioengineering of other peroxygenases to enhance biofuel production. The fact that the mutant P450$_{BSβ}$, in which Q85 is replaced by H85, performs decarboxylation,[6] further emphasizes the role of the LEF in creating function this function.

6. Corresponding Author Details:

1. Prof. Sason Shaik, sason.shaik@gmail.com, sason@yfaat.ch.huji.ac.il, ORCID: 0000-0001-7643-9421

2. Dr. Kshatresh Dutta Dubey, kshatresh@gmail.com, kshatresh.dubey@snu.edu.in, ORCID: 0000-0001-8865-7602

7. Supplementary Materials:

 The Supporting Information is available free of charge on the ACS Publications website at http://pubs.acs.org. It contains the coordinates for the QM region for different reactive species, charge and spin densities.

6. References

1. a) Ortiz de Montellano, P. R. (Ed.) Cytochrome P450: Structure, Mechanism and Biochemistry. 2nd Edition, Plenum Press: New York, **1995**. b) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. Heme-containing oxygenases. *Chem. Rev.* **1996,** *96*, 2841−2888.c) Ortiz de Montellano, P. R. Hydrocarbon hydroxylation by cytochrome

P450 enzymes. *Chem. Rev.* **2010,** *110*, 932−948. d) Groves, J. T. Using push to get pull. *Nat. Chem.* **2014,** *6*, 89−91. e) van Eldik, R. Introduction: inorganic and bioinorganic mechanisms. *Chem. Rev.* **2005,** *105*, 1917−1922. f) Dubey, K. D.; Shaik, S. Cytochrome P450-the wonderful nanomachine revealed through dynamic simulations of the catalytic cycle. *Acc. Chem. Res.* **2019,** *52*, 389−399.

2.   a) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **2013,** *339*, 307-310. b) Sharon, D. A.; Mallick, D.; Wang, B.; Shaik, S. Computation sheds insight into iron porphyrin carbene's electronic structure, formation, and N–H insertion reactivity. *J. Am. Chem. Soc.* **2016,** *138*, 9597-9610.

3.   a) Prier, C. K.; Zhang, R. K.; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme. *Nat. Chem.* **2017,** *9*, 629-634. b) Roiban, G. D.; Reetz, M. T. Enzyme promiscuity: using a P450 enzyme as a carbene transfer catalyst. *Angew. Chem. Int. Ed.* **2013,** *52*, 5439-5440.

4.   a) Meunier, B.; de Visser, S. P.; Shaik, S. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.* **2004,** *104*, 3947-3980. b) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by QM/MM Calculations. *Chem. Rev.* **2010,** *110*, 949-1017.

5.   a) Arnold, F. H. Design by directed evolution. *Acc. Chem. Res.* **1998,** *31*, 125-131. b) Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol*. **2009,** *10*, 866-876.c) Arnold, F. H. Innovation by evolution: Bringing New Chemistry to Life (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019,** *58*, 14420-14426. d) Chen, K.; Arnold, F. H. Engineering new catalytic activities in enzymes. *Nat. Catal.* **2020,** *3*, 203-213.

6.   a) Reetz, M. T. Directed evolution of selective enzymes: catalysts for organic chemistry and biotechnology. *John Wiley & Sons.* **2016.** b) Wang, J. B.; Li, G.; Reetz, M. T. Enzymatic site-selectivity enabled by structure-guided directed evolution. *Chem. Commun.* **2017,** *53*, 3916-3928.c) Ilie, A.; Reetz, M. T. Directed evolution of artificial metalloenzymes. *Isr. J. Chem.* **2015,** *55*, 51-60. d) Wang, J. B.; Lonsdale, R.; Reetz, M. T. Exploring substrate scope and stereoselectivity of P450 peroxygenase OleT JE in

olefin-forming oxidative decarboxylation. *Chem. Commun.* **2016,** *52*, 8131-8133; e) Liu, Y., Wang, C., Yan, J. Zhang, W.; Guan, W.; Lu, X.; Li, S. Hydrogen peroxide-independent production of *α*-alkenes by OleT$_{JE}$ P450 fatty acid decarboxylase. *Biotechnol Biofuels,***2014** *7***,** 28.

7. a) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. *Chem. Rev.* **2005,** *105*, 2279-2328; b) Shaik, S.; Hirao, H.; Kumar, D. Reactivity Patterns of Cytochrome P450 Enzymes: Multifunctionality of the Active Species, and the Two States-Two Oxidants Conundrum. *Nat. Prod. Rep.***2007**, *24*, 533-552.

8. Munro, A. W.; McLean, K. J.; Grant, J. L.; Makris, T. M. Structure and function of the cytochrome P450 peroxygenase enzymes. *Biochem. Soc. Trans.* **2018,** *46*,183-196.

9. Matsunaga, I.; Sumimoto, T.; Ueda, A.; Kusunose, E.; Ichihara, K. Fatty acid-specific, regiospecific, and stereospecific hydroxylation by cytochrome P450 (CYP152B1) from Sphingomonas paucimobilis: Substrate structure required for α-hydroxylation. *Lipids* **2000,** *35*, 365-371.

10. Matsunaga, I.; Ueda, A.; Fujiwara, N.; Sumimoto, T.; Ichihara, K. Characterization of the ybdT gene product of Bacillus subtilis: novel fatty acid β-hydroxylating cytochrome P450. *Lipids* **1999,** *34*, 841-846.

11. Kalita, S.; Shaik, S.; Kisan, H. K.; Dubey, K. D. A Paradigm Shift in the Catalytic Cycle of P450: The Preparatory Choreography during O$_2$ Binding and Origins of the Necessity for Two Protonation Pathways. *ACS Catal.* **2020,** *10*, 11481-11492.

12. Lunine, J. I. Physical conditions on the early Earth. *Phil. Trans. R. Soc. B.* **2006,** *361*, 1721-1731.

13. Lyons, T. W.; Reinhard, C. T.; Planavsky, N. J. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **2014**, *506*, 307-315.

14. Grant, J. L.; Hsieh, C. H.; Makris, T. M. Decarboxylation of fatty acids to terminal alkenes by cytochrome P450 compound I. *J. Am. Chem. Soc.* **2015,** *137*, 4940-4943.

15. Hsieh, C. H.; Huang, X.; Amaya, J. A.; Rutland, C. D.; Keys, C. L.; Groves, J. T.; Austin, R. N.; Makris, T. M. The enigmatic P450 decarboxylase OleT is capable of, but evolved to frustrate, oxygen rebound chemistry. *Biochemistry* **2017,** *56*, 3347-3357.

16. Ramanan, R.; Dubey, K. D.; Wang, B.; Mandal, D.; Shaik, S. Emergence of function in P450-proteins: a combined quantum mechanical/molecular mechanical and molecular dynamics study of the reactive species in the H_2O_2-dependent cytochrome P450 SPα and Its regio-and enantioselective hydroxylation of fatty acids. *J. Am. Chem. Soc.* **2016,** *138*, 6786-6797.

17. a) Belcher, J.; McLean, K. J.; Matthews, S., Woodward, L. S.; Fisher, K.; Rigby, S. E. J.; Munro, A. W. Structure and biochemical properties of the alkene producing cytochrome P450 OleTJE (CYP152L1) from the Jeotgalicoccus sp. 8456 bacterium. *J. Biol. Chem.* **2014,** *289*, 6535-6550; b) Fang, B.; Xu, H.; Liu, Y.;Qi, F.; Zhang, W.; Chen, H.; Wang, C.; Wang, Y.; Yang, W.; Li, S. Mutagenesis and redox partner analysis of P450 fatty acid decarboxylate OleT$_{JE}$. *Sci. Rep.***2017**,*7,*44258.

18. Rude, M. A.; Baron, T. S.; Brubaker, S.; Alibhai, M.; Del Cardayre, S. B.; Schirmer, A. Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from Jeotgalicoccus species. *Appl. Environ. Microbiol.* **2011,** *77*, 1718-1727.

19. a) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; and Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* 2006, *106*, 3210– 3235, b) Schyman, P.; Lai, W.Z. ; Chen, H.; Wang, Y.; Shaik, S. The Directive of the Protein: How does Cytochrome P450 Select the Mechanism of Dopamine Formation? *J. Am. Chem. Soc.* 2011*, 133*, 7977-7984; c) Fried, S. D.; Boxer, S. G. Electric Field and Enzyme Catalysis, *Annu. Rev. Biochem*., **2017**, *86*, 387; d) Warshel, A. Electrostatic Basis of Structure- Function Correlation in Proteins. Acc. *Chem. Res.,* **1981**, *14*, 284; e) For related approaches, see: Dittner, M.; Hartke, B. Globally Optimal Catalytic Fields for a Dield-Alder Reaction. *J. Chem. Phys*., **2020**, *152*, 114106, f) Welborn, V.V.; Head-Gordon, T. Fluctuations of Electric Field in the Active site of the Enzyme Ketosteriod Isomerase. *J. Am. Chem. Soc*., **2019**, *141*, 12487 - 12492.

20. Grant, J. L.; Mitchell, M. E.; Makris, T. M. Catalytic Strategy for Carbon− Carbon Bond Scission by the Cytochrome P450 OleT. *Proc. Natl. Acad. Sci.* **2016**, *113*, 10049-10054.

21. a)Lee, D. S.; Yamada, A.; Sugimoto, H.; Ogura, H.; Ichihara, K.; Adachi, S. I.; Park, S. Y.; Shiro, Y.; Substrate recognition and Molecular Mechanism of Fatty Acid Hydroxylation by Cytochrome P450 by Bacillus subtilis Crystolographic, Spectroscopic and Mutational Studies. *J. Biol. Chem.* **2003***, 261,*9761-9767. b) Matthews, S.; Belcher,

J. D.; Tee, K. L.; Girvan, H. M.; McLean, K. J.; Rigby, S. E. J.; Levy, C. W.; Leys, D.; Parker, D. A.; Blankley, R. T.; Munro, A. W. Catalytic determinants of alkene production by the cytochrome P450 peroxygenase OleTJE. *J. Bio. Chem*. **2017**, *292*, 5128-5143.

22. Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E. Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle *J. Comput. Chem.* **2012**, *33*, 119– 133

23. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem*. **1993**, *97*, 10269−10280.

24. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation *J. Am. Chem. Soc*. **1993**, *115*, 9620−9631.

25. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **1983**, *79*, 926−935.

26. Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. Langevin stabilization of molecular dynamics *J. Chem. Phys.* **2001**, *114*, 2090−2098.

27. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys*. **1984**, *81*, 3684−3690.

28. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints molecular dynamics of n-alkanes. *J. Comput. Phys*. **1977**, *23*, 327−341.

29. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems *J. Chem. Phys.* **1993**, *98*, 10089−10092.

30. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

31. Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll,

M.; Fahmi, A.; Schäfer, A.; Lennartz, C. QUASI: a general-purpose implementation of the QM/MM approach and its application to problem in catalysis. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1−28.

32. Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. Chemshell-a modular software package for QM/MM simulations. *Comput. Mol. Sci.* **2014**, *4*, 101−110.

33. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: the program system turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165−169.

34. Smith, W.; Forester, T. R. DL_POLY_2.0: a general-purpose parallel molecular dynamics simulation package *J. Mol. Graphics* **1996**, *14*, 136−141.

35. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696-371.

36. Becke, A. D. Density Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

37. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction, *J. Comp. Chem.* **2006,** *27*, 1787-99.

38. Kastner, J.; Carr, J.M.; Keal, T.W.; Thiel, Wander, W. A.; Sherwood, P. DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations. *J. Phys. Chem. A*, **2009**, *113*, 11856–11865.

39. a) Stuyver, T.; Huang, J.; Mallick, D.; Danovich, D.; Shaik, S. TITAN: A Code for Modeling and Generating Electric Fields-Features and Applications to Enzymatic Reactivity. *J. Comput. Chem*. **2020**, *41*, 74-82; b) Stuyver, T.; Danovich, D.; Shaik, S. "CHAPTER 6: Computational Generation and Quantification of Electric Fields and Electrostatic-mediated Catalyst Optimization" In "Electric Fields and Structure-Reactivity Aspects: New Horizons in Chemistry" *R. Soc. Chem.* **2021,** 195-224.

40. Balasubramani, S. G.; Chen, G. P.; Coriani, S.; Diedenhofen, M.; Frank, M. S.; Franzke, Y. J.; Furche, F.; Grotjahn, R.; Harding, M. E.; Hättig, C.; Hellweg, A.; Helmich-Paris, B.; Holzer, C.; Huniar, U.; Kaupp, M.; Khah, A. M.; Khani, S. K.; Müller, T.; Mack, F.; Nguyen, B. D.; Parker, S. M.; Perit, E.; Rappoport, D.; Reiter, K.; Roy, S.; Rückert, M.; Schmitz, G.; Sierka, M.; Tapavicza, E.; Tew, D. P.; van Wüllen, C.; Voora, V. K.;

Weigend, F.; Wodyński, A.; Yu, J. M. TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020,** *152*, 184107.

41. a) Shaik, S.; Stuver, T. CHAPTER 1: The Introduction of the Book, In: *Electric Fields and Structure-Reactivity Aspects: New Horizons in Chemistry*, *R. Soc. Chem.,* London, **2021,** 1-11, Shaik, S.; Stuyver, T. Eds.; b) Shaik, S.; Danovich, D.; Dubey, K. D.; Stuyver, T. CHAPTER 2: The Impact of Electric Fields on Structure and Reactivity, In: *Electric Fields and Structure-Reactivity Aspects: New Horizons in Chemistry*, *R. Soc. Chem.* **2021,** 12-70, Shaik, S.; Stuyver, T. Eds.; c) Blyth, M. T.; Coote, M. L. CHAPTER 4: *Recent advances in designed local electric fields*, In *Effects of Electric Fields on Structure and Reactivity: New Horizons in Chemistry*, Shaik, S.; Stuyver, T., Eds. *R. Soc. Chem.* **2021,** 119-146; d) Sowlati-Hashjin, S.; Karttunen, N.; Matta, C.F. Electrostatic Fields in biophysical chemistry. In: *Electric Fields and Structure-Reactivity Aspects: New Horizons in Chemistry*, *R. Soc. Chem.,* London, **2021,** 225-262, Shaik, S.; Stuyver, T. Eds.

42. Faponle, A. S.; Quesne, M. G.; de Visser, S. P. Origin of the Regioselective Fatty-Acid Hydroxylation versus Decarboxylation by a Cytochrome P450 Peroxygenase: What Drives the Reaction to Biofuel Production? *Chem. Eur. J.* **2016**, *22*, 5478.

43. a) Bell, R. P. The Theory of Reactions involving Proton Transfer. *Proc. R. Soc. London, Ser. A,* **1936**, *154*, 414. b) Evans, M. G.; Polanyi, M. Further Considerations on the Thermodynamics of Chemical Equilibria and Reaction Rate *J. Chem. Soc., Faraday Trans.*, **1936**, *32*, 1333.

Table of content (TOC)